AGU PUBLICATIONS

JGR

Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE

10.1002/2015JD024582

Key Points:
- Three new daily instrumental series dating back to the late nineteenth century are presented
- Cloud cover is a valuable constraint for the homogenization of daily temperature series
- Summer hot days in the southern Alps have increased by over 200% in the last 140 years

Supporting Information:
- Supporting Information S1

Correspondence to:
Y. Brugnara,
yuri.brugnara@giub.unibe.ch

Citation:
Brugnara, Y., R. Auchmann, S. Brönnimann, A. Bozzo, D. C. Berro, and L. Mercalli (2016), Trends of mean and extreme temperature indices since 1874 at low-elevation sites in the southern Alps, *J. Geophys. Res. Atmos.*, *121*, 3304–3325, doi:10.1002/2015JD024582.

Received 2 DEC 2015
Accepted 20 MAR 2016
Accepted article online 25 MAR 2016
Published online 8 APR 2016

Trends of mean and extreme temperature indices since 1874 at low-elevation sites in the southern Alps

Yuri Brugnara[1,2], Renate Auchmann[1,2], Stefan Brönnimann[1,2], Alessio Bozzo[3], Daniele Cat Berro[4], and Luca Mercalli[4]

[1]Institute of Geography, University of Bern, Bern, Switzerland, [2]Oeschger Centre for Climate Change Research, Bern, Switzerland, [3]European Centre for Medium-Range Weather Forecasts, Reading, UK, [4]Società Meteorologica Italiana, Torino, Italy

Abstract We describe the recovery of three daily meteorological records for the southern Alps (Domodossola, Riva del Garda, and Rovereto), all starting in the second half of the nineteenth century. We use these new data, along with additional records, to study regional changes in the mean temperature and extreme indices of heat waves and cold spells frequency and duration over the period 1874–2015. The records are homogenized using subdaily cloud cover observations as a constraint for the statistical model, an approach that has never been applied before in the literature. A case study based on a record of parallel observations between a traditional meteorological window and a modern screen shows that the use of cloud cover can reduce the root-mean-square error of the homogenization by up to 30% in comparison to an unaided statistical correction. We find that mean temperature in the southern Alps has increased by 1.4°C per century over the analyzed period, with larger increases in daily minimum temperatures than maximum temperatures. The number of hot days in summer has more than tripled, and a similar increase is observed in duration of heat waves. Cold days in winter have dropped at a similar rate. These trends are mainly caused by climate change over the last few decades.

1. Introduction

The Alpine region is one of the areas in Europe most vulnerable to climate change [*European Environment Agency*, 2009]. It has already experienced an extraordinary rise in temperature since the late nineteenth century, at more than twice the rate of the global trend [*Auer et al.*, 2007]. This change is confirmed by the rapid retreat of the Alpine glaciers, amounting to an average loss of 2 Gt yr^{-1} of ice over the period 2003–2009 [*Gardner et al.*, 2013].

The Alps represent a fascinating mixture of different climates, because of the vertical extension and geographical position. In particular, the Alpine ridge strongly interacts with the midlatitude storm track, which results in important climatological differences between the southern and northern sides of the chain [*Frei and Schär*, 1998; *Auer et al.*, 2007; *Brunetti et al.*, 2009]. The presence of the Mediterranean Sea to the south further enhances these differences. On a more local scale, the complex topography and numerous lakes allow for a multitude of microclimatic peculiarities.

The change in mean temperature in the Alpine region over the last two centuries has been thoroughly studied thanks to the availability of long monthly instrumental series [*Auer et al.*, 2007; *Brunetti et al.*, 2009]. Trends in temperature extremes still lack such solid results, particularly in the southern Alps. This is due to the limited recovery of long daily records, as well as the difficulties in homogenizing daily data. The most significant work of this kind was that of *Della-Marta et al.* [2007a], who analyzed changes in summer heat waves in western Europe using 44 homogenized daily temperature series covering the period 1880–2005. They found that the length of summer heat waves doubled during this period and that the frequency of hot days almost tripled. While 10 stations used by *Della-Marta et al.* [2007a] are located in the northern Alps, only one (Lugano, Switzerland) is from the southern Alps. Our goal is to extend this analysis to the southern Alps by using multiple stations; moreover, we extend the analysis to cold extremes in winter.

Climate data series are usually affected by nonclimatic signals (inhomogeneities) caused by station relocations, changes in the instruments and the radiation screen, changes in the measurement procedures, etc. In order to analyze real climatic trends, these inhomogeneities must be detected and removed, a process called homogenization. Compared to annual or monthly data, inhomogeneities in daily data can affect moments of the distribution other than the mean (e.g., variance). The correction of high-order moments in daily data is particularly important for the study of extremes, but this is also a challenging task [*Brunet et al.*, 2008].

In the literature, two approaches have been proposed to homogenize high-order moments of daily temperature: a fully statistical approach [e.g., *Della-Marta and Wanner*, 2006; *Toreti et al.*, 2010; *Mestre et al.*, 2011; *Trewin*, 2013] and a physics-based approach that directly addresses the causes of the inhomogeneity [e.g., *Auchmann and Brönnimann*, 2012]. The latter requires extremely detailed metadata, which are usually not available, and therefore, this method is difficult to apply. However, a third possible solution is to combine the two approaches by adapting the statistical procedure, taking into account covariates in the estimation process. In particular, *Auchmann and Brönnimann* [2012] considered radiation (modeled using cloud cover observations), wind speed, and ground albedo as the fundamental physical factors influencing the day-to-day variability of an inhomogeneity that depends on the change of the radiation screen. Such relations were also found by analyzing parallel observations at different sites in Europe [*Brunet et al.*, 2008; *Van der Meulen and Brandsma*, 2008; *Huwald et al.*, 2009]. For instance, the radiation screen usually has a negligible impact on the observed temperature in very windy and cloudy conditions but can cause large biases in comparison with other types of screens during a still, sunny day.

Up until at least the midtwentieth century, a wide variety of radiation screens were used at meteorological observatories [see e.g., *Parker*, 1994], causing significant biases in the temperature observations. Even modern screens can produce biases in excess of 1°C, when cloud cover and wind speed are low [e.g., *Van der Meulen and Brandsma*, 2008]. Similar considerations apply to the position of the screen (e.g., distance from the ground and other heat sources). One can therefore expect that cloud cover may be a valuable constraint for the correction of inhomogeneities caused by screen changes and station relocations. In this work, we use subdaily cloud cover observations as an aid for the statistical model in the estimation of the corrections. We are not aware of any previous study that has attempted to use cloud cover for a statistical homogenization of historical daily temperature records, probably because of the lack of digitized cloud cover data. We investigate the potential benefit of using cloud cover observations by comparing a parallel record of a classical meteorological window and a modern screen.

This paper is organized into five main sections. Section 2 describes three unpublished instrumental records for the southern Alps that we recovered. Section 3 describes the methods, in particular, those that we use to address data quality and homogeneity. In section 4, we evaluate the homogenization algorithm in the parallel record before we analyze trends in the mean and extreme indices, which are then discussed in section 5. Section 6 contains our concluding remarks.

2. Data

2.1. Temperature

Three records of daily maximum (T_x) and minimum (T_n) temperature (Domodossola, Riva, and Rovereto) were independently recovered and digitized by the authors over the past few years, along with the relevant metadata (observation times, observers' names, exact station location, type and position of instruments, and their changes over time). Monthly averages for these sites had already been recovered and analyzed in previous studies [e.g., *Brunetti et al.*, 2006; *Auer et al.*, 2007], although they are not necessarily consistent with our data due to differences in the data sources. For example, the mentioned studies used a different station in Rovereto for the recent years (M. Brunetti, personal communication, 2016).

We also use the daily record of Lugano, which was already analyzed by *Della-Marta et al.* [2007a]. However, we start from the original raw data and perform a new homogenization.

Figure 1 shows the geographical position of the stations (red points), which are clustered in two pairs of very close locations. The stations are located at low elevations (< 300 m) at the bottom of wide valleys, thus presenting similar climatological characteristics. The climates of Lugano and Riva, however, are influenced by the presence of large lakes.

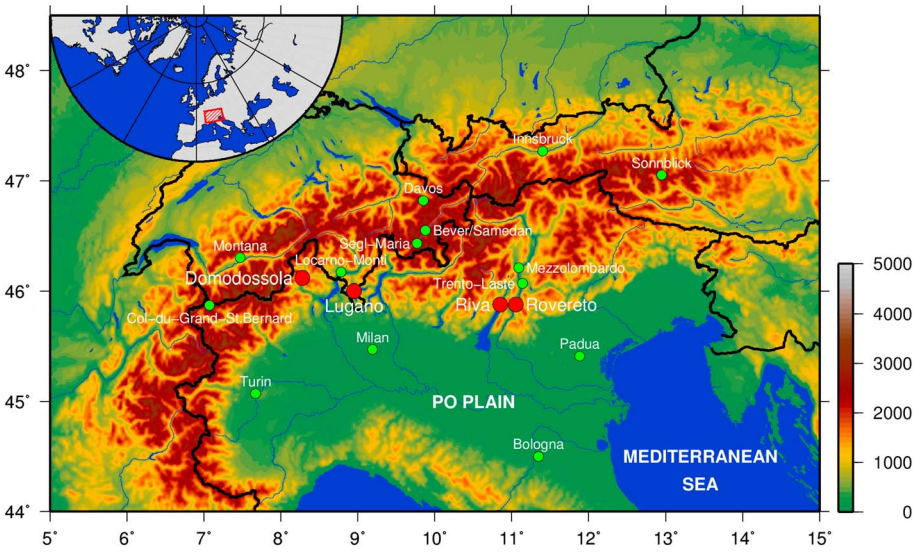

Figure 1. Map of the stations with topography (in meters above mean sea level). The stations used for trend analysis are indicated by the red points.

2.2. Cloud Cover

We also digitized subdaily cloud cover observations, representing the fraction of the sky obscured by clouds. They are expressed either in tenths (Domodossola, Riva, and Rovereto) or eighths/octas (Lugano). A value of 0/10 (0/8) indicates the absence of clouds, while a 10/10 (8/8) represents an entirely covered sky. The data that we use are visual observations made by human observers. The observations were performed 3 times per day at fixed hours (slightly different for each station), but we only use the two observations made in the morning and the afternoon, which are usually close to the times when daily minimum and maximum temperatures, respectively, occur.

The replacement of an observer can have an impact on the homogeneity of the observations over time, mainly because the observers (in most cases clergymen) were not always sufficiently instructed. This is relevant for a trend analysis of cloud cover, but less important for our purposes.

2.3. Domodossola

The record of Domodossola was recovered by the Italian Meteorological Society (SMI) from the original meteorological registers (1871–1972), conserved at the *Collegio Rosmini* in Domodossola. Unfortunately, a few years (1932–1937, 1944, 1945, 1950, 1961–1962, and 1973–1987) were missing. From 1988 onward daily data were provided by different institutions (Mountain Community Valle Ossola, Board for the Agricultural Development of Piedmont, and the National Research Council) that had progressively taken over the management of the station.

Domodossola lies in the Ossola Valley in the north of the Italian region of Piedmont, at the confluence of five secondary valleys. Its elevation is approximately 280 m above mean sea level (amsl).

The meteorological observatory of Domodossola was inaugurated on 30 November 1871, by initiative of the Italian Alpine Club in collaboration with the SMI. The instruments were installed in the town center ("a" in Figure S1 in the supporting information). The local Rosminian monks were responsible for the observations for over a century, until 1987. The station had a director who functioned as the scientific supervisor; some of the directors were also active in publishing studies based on the data measured in Domodossola [e.g., *Pinauda*, 1914; *Pattarone and Alice*, 1925].

The standard observation times were fixed at 9:00, 15:00, and 21:00 local time (LT). In June 1876, the station was relocated to a new building, approximately 250 m from the old one ("b" in Figure S1), but the exposure of the thermometers was kept nearly unchanged (i.e., meteorological window on a tower). In February 1905, the thermometers were moved from the tower to a freestanding wooden screen in the backyard ("c" in Figure S1).

2.4. Riva

The record of Riva was partly (1874–1915) recovered at the University of Bern from the meteorological yearbooks of the former Austro-Hungarian Empire. The second part of the series (1949–1974) was also transcribed at the University of Bern from data sheets stored in the archive of the former Italian Central Office of Agrometeorology, which were recently photographed and published by a consortium of Italian research institutes (project Recupero e valorizzazione dell'Archivio meteo Storico TRentinO (ASTRO)). The third part (1975–2015) is available on the website of the regional weather service of the Autonomous Province of Trento (Meteotrentino).

Riva (also known as Riva del Garda) is located on the northern tip of Lake Garda, the largest lake of Italy (Figure 1). Riva's climate is largely influenced by the lake [e.g., *Giovannini et al.*, 2015], which mitigates temperatures particularly near the shore, where the town center lies (Figure S2). The elevation of Lake Garda is 65 m amsl.

Until the end of World War I, Riva was part of the Austro-Hungarian Empire. Official meteorological observations in Riva began in 1869, but complete daily observations have survived only from 1874, when they began to be published in extenso in the yearbooks of the *Central-Anstalt für Meteorologie und Erdmagnetismus*. The yearbooks also reported relocations and detailed descriptions of the stations based on inspections by officials of the Central-Anstalt. Temperature and cloud cover, along with other variables, were observed 3 times per day at 7:00, 14:00, and 21:00 LT. T_x and T_n measurements are available starting from July 1876, when maximum and minimum thermometers were added to the instrumentation. The director of the local gymnasium, Dario Bertolini, took care of the station (which was moved to his house in 1888) for over 40 years (1871–1912).

In May 1915, Italy declared war on Austria-Hungary. The town of Riva, located on the front line, was entirely evacuated. The interruption in the temperature observations lasted until 1925, when the Italian Hydrographic Service reactivated the station. Unfortunately, only the data after 1948 could be found at daily resolution.

The station has undergone several relocations, the most significant being the move from the town center to a hydroelectric power plant located in a rural area approximately 1.5 km from the lake (1971; "d" in Figure S2). Recently (2012), the station was relocated to the village of Torbole (3 km southeast of Riva, "e" in Figure S2).

For the period 1978–1990, we shifted backward the T_x observations by 1 day. This is because, following the standard of the Italian Hydrographic Office, the observations (made in the morning) during that period were assigned to the calendar day in which the observations were made, even though the maximum value was usually reached in the previous calendar day.

2.5. Rovereto

Rovereto is the second largest town in the Italian province Trentino. Similar to Riva, it was part of the Austro-Hungarian Empire until 1919. The town lies about 15 km east of Riva (Figure 1) in one of the major Alpine valleys, oriented in the north-south direction along the river Adige.

Weather observations in Rovereto were sporadic until the opening of the meteorological observatory in January 1882 by the joint effort of the *Società Alpinistica Tridentina*, the Natural History Museum of Rovereto, and the Austrian Central-Anstalt, with the collaboration of the SMI. The cloister of the Franciscan monks was chosen to host the observatory, with the instruments housed inside a wooden case (Figure 2) on the north facing side of the building about 5 m above the ground, following the common guidelines in those times [*Jelinek*, 1869; *Denza*, 1882]. The cloister is located in the northern part of the town at an elevation of 207 m amsl.

The weather records suffered only two major interruptions, both linked to the wars that occurred in Europe in the first half of the twentieth century. The longest break occurred between 1915 and 1919, when Rovereto was evacuated and severely damaged. The Italian Army restored the station in the same building immediately after the end of the war. A second interruption occurred in the autumn of 1943 during World War II.

In Rovereto, observations have always been made at the same site. Minor movements along the same north facing wall are recorded in 1919 and 1960, when the building was renovated. In the early 1990s, the Italian Hydrographic Service and Meteotrentino set up a second automatic measurement site within the cloister grounds, providing WMO-compliant observations parallel to the historic site (Figure 2). This automatic weather station has undergone several interruptions since 2004 due to problems with instrumentation.

Figure 2. Pictures of the meteorological observatory of Rovereto taken around 2003: (top) the meteorological window and its sorroundings, (bottom) the automatic weather station (arrow indicates the thermometer), and (right) the inside of the meteorological window.

The standard observation times of Rovereto were 9:00, 15:00, and 21:00 LT, until December 1932, when they changed to the current times of 8:00, 14:00, and 19:00 LT. The data were recovered directly from the original registers conserved by the Natural History Museum of Rovereto.

2.6. Lugano and Other Reference Series

The record of Lugano was provided by the Swiss National Weather Service (Meteoswiss). Metadata come from *Kuglitsch et al.* [2012]. This series is the longest record available for this study, covering the period 1864–2015 without gaps.

Lugano is located on the northwestern shore of lake Lugano in the Swiss canton of Ticino, approximately 50 km southeast of Domodossola (Figure 1). Similarly to Riva, its climate is mitigated by the proximity of the lake. The station has always been located near the shore, despite two relocations in 1940 and 1972. The elevation of the lake is 271 m amsl. Observation times were at 7:30, 13:30, and 21:30 LT until 1970 and at 6:45, 12:45, and 18:45 LT from 1971 to 1980. Automatic instruments were introduced in 1981, allowing temperature measurements every 10 min. Since then, daily extremes refer to the UTC calendar day (i.e., starting at 1:00 LT).

We use 14 additional T_x and T_n series (Table 1) as references for the homogenization. Six of these come from Meteoswiss [*Begert et al.*, 2005; *Kuglitsch et al.*, 2012], five are from the European Climate Assessment and Dataset [*Klein Tank et al.*, 2002], two are from Meteotrentino, and one was provided by the SMI [*Di Napoli and Mercalli*, 2008]. Their geographical positions are shown in Figure 1. For 11 of these series (see Table 1) we only use annual and semiannual means, because they are only used for the detection of inhomogeneities and not for the calculation of the corrections (see section 3.3).

3. Methods

3.1. Quality Control

Most data series used in this paper (excluding those digitized by the authors) have undergone a quality control by the respective data providers. Nevertheless, for the sake of consistency we performed quality tests on each of the four series used in the analysis, as well as on the three additional series (Locarno-Monti, Mezzolombardo, and Trento-Laste) used as references for the calculation of corrections (see section 3.3). First, we carried out a gross quality test on daily data based on absolute thresholds for unrealistic values (seasonally differentiated; e.g., $T_x > 35^\circ$C in winter and $T_x > 45^\circ$C in summer). We then compared the daily observations in each station with the average of the observations in the other six stations for the same day. A difference larger

Table 1. Additional Temperature Series Used as Reference for the Break Detection[a]

Station[b]	Lon	Lat	Elev	Period	Gaps	Source	Reference[c]
Bever/Samedan	9.88	46.55	1710	1869–2015	–	Meteoswiss	–
Bologna	11.35	44.50	53	1814–2003	–	ECA and D	–
Col du Grand-St.Bernard	7.07	45.87	2472	1864–2015	1901–1963	Meteoswiss	–
Davos	9.85	46.82	1590	1876–2015	1883–1888	Meteoswiss	–
Innsbruck	11.40	47.27	577	1877–2015	–	ECA and D	*Nemec et al.* [2013]
Locarno-Monti	8.78	46.17	366	1935–2015	–	Meteoswiss	–
Mezzolombardo	11.09	46.21	225	1925–2006	–	Meteotrentino	–
Milan	9.19	45.47	132	1863–1998	–	ECA and D	*Maugeri et al.* [2002]
Montana	7.47	46.30	1427	1931–2015	–	Meteoswiss	–
Padua	11.88	45.41	31	1854–1997	–	ECA and D	*Cocheo and Camuffo* [2002]
Segl-Maria	9.77	46.43	1798	1869–2015	1915–1976	Meteoswiss	–
Sonnblick	12.95	47.05	3106	1886–2015	–	ECA and D	*Nemec et al.* [2013]
Trento-Laste	11.14	46.07	312	1920–2015	–	Meteotrentino	–
Turin	7.67	45.07	277	1753–2009	–	SMI	*Di Napoli and Mercalli* [2008]

[a]Lon = longitude in degrees East, Lat = latitude in degrees North, Elev = elevation in meters, Period = period covered by the record, Gaps = large data gaps, Source = data provider.
[b]The names of the stations that are also used for the calculation of the corrections are written in italic.
[c]The records without reference are not homogenized.

than 15°C for T_x and 10°C for T_n was considered suspicious and was subjected to a more detailed inspection. Of the 147 daily values flagged as erroneous within the seven series, 89 originated from digitization errors or mistranscriptions and could be corrected by consulting the original sources. The remaining values, mostly caused by malfunction of the modern electronic sensors, were excluded from the analysis.

Finally, we inspected seasonal averages and their differences from other stations in order to detect irregular or short-lasting inhomogeneities (due, e.g., to a malfunctioning instrument) for which a reliable estimation of corrections is not possible. This led to the removal of a relevant fraction of data in the series of Domodossola, in particular, a long period (July 1876 to June 1885) after the first relocation of the station (from a to b in Figure S1), when large problems with T_x appear (interannual variability was too high). The maximum thermometer was probably damaged during the relocation; in fact, T_x observations are missing for the first 20 days of July 1885 suggesting that the thermometer was sent for repair. Incoherent data were also evident for short periods in 1889 (T_n only), between 1934 and 1935, and in 1988 (T_x only). Moreover, we removed all data for August and September 1886 because, according to the metadata, the station was temporarily moved into an attic to allow for renovation works during those months. In the other series, we removed incoherent T_n observations during winter 1957/1958 for Rovereto and during parts of 1928 and 1929 for Lugano, as well as some data in the years 1912 to 1915, 1952, 1953, and 1965 for Riva.

Figure 3 shows the temporal coverage of the series and data gaps after the quality control.

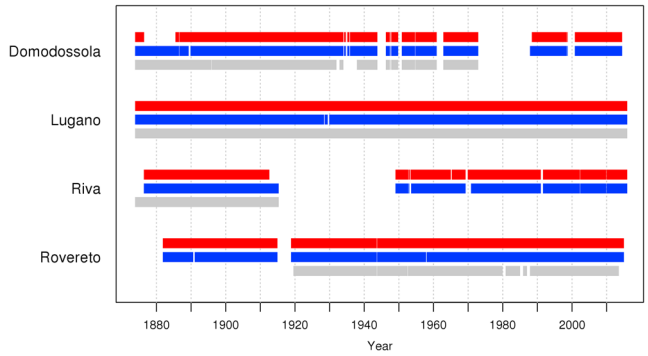

Figure 3. Timeline of data availability for daily T_x (red), T_n (blue), and subdaily cloud cover (grey) over the analyzed period (1874–2015).

3.2. Observation Times

Changes of the observation times represent a problem affecting nearly all historical records. In particular, the introduction of automatic instruments implied a worldwide redefinition of the meteorological day, which nowadays usually coincides with the calendar day (i.e., starting at midnight). Such changes are a potential source of inhomogeneity for daily extremes. *Trewin* [2013], for instance, found a shift of a few tenths of degrees Celsius in the annual mean of T_n in many Australian stations, caused by the adoption of the calendar day.

One of the stations analyzed in our study, Riva, is affected by several changes of the observation time: both T_x and T_n readings were made at 21:00 LT during 1876–1915, 17:00 LT during 1949–1969, 7:00 LT in 1970, 9:00 LT during 1971–1990, and 0:00 LT during 1991–2015. We tested whether these changes cause shifts in the mean values by using temperature observations made in Riva (power plant, d in Figure S2) every 15 min between 1 July 2009 and 30 June 2012. In agreement with the results of *Trewin* [2013], we found an average systematic difference of −0.16°C for T_n for the transition from 9:00 LT to 0:00 LT. This transition coincides with a significant instrumentation change; therefore, we did not apply a specific correction for the observation time. The inhomogeneity as a whole is corrected at a later stage by the homogenization algorithm (section 3.3). We also calculated the impact of the observation time on the number of T_n observations below −1°C (roughly corresponding to the 10th percentile in winter) and T_x observations above 33°C (approximately 90th percentile in summer). The 7:00 LT observation time causes a relatively large overestimation (approximately +10%) in the count for T_n, so we decided not to use the T_n observations of the year 1970 in the analysis. All other observation times give a count of 48 ± 1 days. The count for T_x is less affected by the observation time; the largest differences amount to only 1 day (approximately 2%). Note that before 1971, however, the station was located in the town center, where we do not have subhourly observations. Shadows from the surrounding mountains and other microclimatic features, as well as the characteristics of the radiation screen, could influence the impact of the observation time on the daily extremes.

A similar analysis for the station of Lugano, using observations made every 10 min during the period 1981–2010, revealed an average difference of −0.26°C for T_n caused by the change of observation time from 18:45 LT to 1:00 LT in 1981. However, with the introduction of the automatic network, the Swiss weather service changed the definition of daily extremes so that T_n is defined as the minimum of the instantaneous temperature observations registered every 10 min, causing a mean overestimation of T_n of arguably the same magnitude as the underestimation introduced by the change of observation time. Given the impossibility to verify this with the data in our possession, we did not apply any correction at this stage.

Observation times for the daily extremes at the station of Domodossola are uncertain. In the early years the station probably followed the guidelines of the SMI [*Denza*, 1882], which required the reading of T_n at 9:00 LT and of T_x at 21:00 LT. This practice, however, might have changed during the twentieth century. The calendar day was adopted in 1988, when we found a significant inhomogeneity in both variables (see section 3.3.1) that is corrected by the homogenization algorithm.

The changes of observation times at the station of Rovereto were small and are not expected to have a significant impact on daily extremes. In particular, that station has not yet adopted the calendar day.

3.3. Homogenization

3.3.1. Break Detection

The first step of the homogenization procedure is the definition of homogeneous subperiods in the series, delimited by so-called "breakpoints," which will then be used for the calculation and application of the corrections.

Our breakpoint detection procedure (hereafter "break detection") is an adaptation of the approach described in *Kuglitsch et al.* [2012]; we combine three different statistical methods, each one making use of reference series at annual temporal resolution. The methods, described by *Caussinus and Mestre* [2004], *Wang et al.* [2007], and *Toreti et al.* [2012], respectively, use automatic algorithms based on different segmentation techniques applied to a series of differences (candidate minus reference series). For each combination of method and reference series, we obtain a list of years where breakpoints are detected (note that a particular breakpoint can be caused either by the candidate or the reference series or by both). We define a breakpoint to be significant in the year x if an inhomogeneity is detected in the years from $x - 1$ to $x + 1$ by at least three reference series with at least two of the methods (the reference series that detect the breakpoint can differ between the methods). This approach guarantees a very low rate of false positives [*Kuglitsch et al.*, 2012].

The break detection was performed separately on the annual and "seasonal" (winter season defined from October to March and summer season defined from April to September) averages of T_x, T_n, and the Daily Temperature Range ($DTR = T_x - T_n$). In many cases, we used metadata or particular features of the series (e.g., large gaps) to correct the date of the breakpoint (note that the automatic detection has by definition an uncertainty of ± 1 year). Unless the day of the breakpoint was available from metadata, we assigned the breakpoint to the first day of the year or to the first day after a long interruption in the observations. The analysis of *DTR* did not result in additional breakpoints but was sometimes useful in adjusting the year of breakpoints that were not supported by metadata, in particular, through a visual inspection of the differences with the reference series.

For each candidate series we used the 10 best correlated series as reference series and repeated the break detection on subperiods (>50 years) when one or more reference series have large gaps. These were replaced with others, so that 10 reference series are actually available for each subperiod. This was clearly not possible in the 1860s and part of the 1870s, since few series begin before 1874. The break detection is therefore less reliable before that year. Moreover, 1874 is also the year following the First International Meteorological Congress, held in Vienna in September 1873, which produced the first set of international standards for the observation of meteorological parameters [*Edwards*, 2004]. These considerations led to the selection of the period 1874–2015 for our analysis.

We provide the complete list of detected breakpoints and the causes suggested by metadata in Table S1 in the supporting information. The average homogeneous subperiod was 17 years for both T_x and T_n. About one third of the breakpoints are not supported by metadata. In comparison, *Kuglitsch et al.* [2012] found an average homogeneous subperiod of 48 years in Switzerland, with 94% of the breakpoints supported by metadata. Using a different detection technique applied to a large number of stations (131), *Auer et al.* [2007] found an average homogeneous subperiod of 23 years for mean temperature series in the Alpine region.

3.3.2. Correction

We define h_1 and h_2 as the two homogeneous subperiods before and after a breakpoint in the candidate series (h_2 being the more recent period). Moreover, we define three categories of cloud cover: clear (C_1) for a fraction of covered sky lower than 4/10 (or 3/8, depending on available resolution), cloudy (C_2) for a fraction between 4/10 and 9/10 (3/8–7/8), and overcast (C_3) for a fraction of 10/10 (8/8. This specific definition of the categories is motivated not only by statistical considerations concerning sample size and the conversion between tenths and octas but also by physical properties of clouds. In particular, C_1 conditions are usually associated with *cirrus* and *cumulus* clouds, while C_2 and C_3 have a larger occurrence of *stratus* and *cumulonimbus* types [*Wacker et al.*, 2015]. The category C_3 represents, by definition, an entirely covered sky, meaning that the radiation screen is constantly shaded by the clouds.

In the simplest case of a breakpoint on 1 January and only one reference series, we estimate the corrections $\delta_k(j)$ for each Julian day j ($j = 1, \ldots, 365$) and each cloud cover category C_k ($k = 1, 2, 3$) using a 31 day moving window centered on j. We call $X_c(i, j, k)$ the subsample of temperature observations at the candidate station on days with cloud cover in the category C_k contained in the subperiod h_i ($i = 1, 2$) and within 15 days from the Julian day j, while $X_r(i, j, k)$ is the same for the reference station. The temperature observations on a certain date must be available at both stations; otherwise, they are not used for the calculation of the corrections.

The corrections for the category C_k, when k is either 1 or 3 (clear or overcast), are obtained using the following equation:

$$\delta_k(j) = \left(\overline{X_c(2,j,k)} - \overline{X_r(2,j,k)}\right) - \left(\overline{X_c(1,j,k)} - \overline{X_r(1,j,k)}\right). \quad (1)$$

The corrections are then smoothed over j using a LOcal regrESSion (LOESS) of second order with smoothing parameter $\alpha = 0.25$ [*Cleveland and Devlin*, 1988] and are finally applied to the days with cloud cover in C_k in the subperiod h_1.

For C_2, we use the same procedure without considering the variable k. In other words, to calculate the corrections for this category, we use all days, independent of cloud cover. This choice is motivated by the insufficient size of the samples that would often result from using only the days belonging to C_2, particularly in winter, and from the large spatial and temporal variability that usually characterizes this category, as we will show later on.

In most cases we used more than one reference series and took the weighted average of the corrections calculated from each reference series. The weights are the averages of the linear correlation coefficients between the candidate and reference series in the subperiods h_1 and h_2. Correlations were calculated from daily anomalies [*Trewin*, 2013], using the period 1986–2015 as the reference for the anomalies. When the date of the breakpoint is not 1 January, the days before and after that date in the year of the inhomogeneity are also considered part of h_1 and h_2, respectively, and are used for the estimation of corrections. The homogenization is performed starting from the most recent breakpoint, so that the series becomes homogeneous with respect to the latest observations. T_x and T_n are homogenized separately, using the respective breakpoints.

In addition to the four records that are the target of the homogenization, we used three additional reference series for the calculation of the corrections; these are the shorter records of Locarno-Monti, Mezzolombardo, and Trento-Laste, which are highly correlated ($r > 0.8$) with at least two of the target series. The breakpoints found in these series are also listed in Table S1.

In order to be used for a given breakpoint, a reference series must have homogeneous data for at least 2 years preceding and following the breakpoint, totaling a minimum of 4 years of available observations. One exception to this rule was necessary for the 1988 breakpoint in Domodossola (T_x only), which follows a 15 year long gap; in this case only the year 1972 from the highly correlated reference series of Locarno-Monti ($r = 0.89$) could be used as h_1. We did not set a maximum amount of years to be used; however, these are usually limited by inhomogeneities in the candidate (for h_1 only) and/or in the reference series (note that we always use uncorrected reference series). The average length of h_1 and h_2 across all series and all breakpoints was 10 and 17 years, respectively (including data gaps); h_2 is on average longer because it is only limited by inhomogeneities in the reference series as a consequence of the reversed homogenization.

The final step of our homogenization procedure is to reapply the break detection to the corrected series, in order to verify that all significant breakpoints have been properly corrected. This led, for a few breakpoints, to the manual exclusion of reference series that produced inaccurate corrections.

Ideally, one should estimate the corrections for a certain cloud cover category from the observations made when the same cloud cover is present at the same time at both the candidate and the reference station. In practice, however, this is difficult to apply for two reasons: (1) the number of complete cloud cover series is limited and (2) the sample may become too small for a reliable estimation of the corrections. For these reasons we use only the cloud cover observations of the candidate station (or of the nearest reference station, if not available for the candidate) to define sky conditions at both stations, on the assumption that spatial variability of cloud cover is usually low.

We use the afternoon cloud cover observations for the homogenization of T_x and the morning observations for T_n, since these are expected to be closer to the time of the day when extremes are usually reached. From a physical point of view, cloud cover in daytime affects near-surface temperature observations mainly by modulating incoming shortwave radiation, whereas the energy budget during night time is strongly affected by longwave radiation emitted by the clouds [*Auchmann and Brönnimann*, 2012].

Despite the intrinsic subjectivity of visual observations and despite possible differences in the observation procedure between different countries, the climatological distribution over the three categories C_k is very similar within the analyzed stations (Table S2). The observations of clear sky represent more than 42% of the total, while those of overcast sky vary between 23 and 30%, depending on the station.

Similarity between cloud cover at each station is the main assumption of our method and needs to be verified. Figure 4 shows the aggregate distribution of "simultaneous" (i.e., during the same part of the day, not necessarily at exactly the same time) cloud cover observations at three of the four stations used in the analysis (references) when a certain category is observed at the station that is left out (candidate). When the category at the candidate station is either C_1 or C_3 (Figure 4 top and bottom rows), then the category at any of the reference stations coincides in more than 60% of the cases. The percentage is lower for cloudy days (C_2; Figure 4, middle row) because they are, in general, associated with more variable weather conditions. The probability of observing a certain category at a reference station, when the category C_2 is observed at the candidate station, is not much different from the climatological probability (Table S2). Hence, using all days to calculate the corrections for C_2, independent of their cloud cover, will give similar results to using cloudy days only, with the advantage of having a larger sample size.

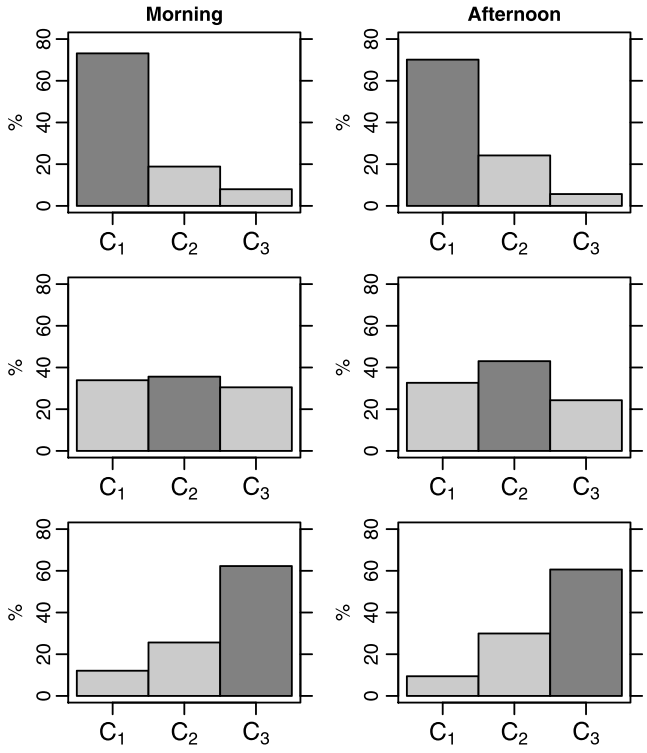

Figure 4. Fractions of cloud cover observations at the reference stations falling into the three categories C_k, when a certain category (indicated by the dark bar) is observed at the candidate station for (left column) morning and (right column) afternoon observations. For example, (top row) shows the fraction of clear, cloudy, and overcast skies observed at the reference stations when the sky is clear at the candidate station. Seasonal versions of this figure are provided in the supporting information.

The influence of the geographical distance on the similarity between cloud cover conditions at different stations is relatively small, although the probability of large differences (clear at the candidate station and overcast at the reference station or vice versa) is much lower when the stations are close to each other (Figures S3 and S4).

These results demonstrate that the spatial and temporal variability of cloud cover in the southern Alps is relatively small when cloud cover belongs to either C_1 or C_3. These climatic conditions are favorable for using this parameter as a homogenization aid. However, the variability is arguably higher when comparing stations at vastly different altitudes because of low clouds in winter and orographically induced convection in summer.

3.3.3. Parallel Record

We use 10 years (1992–2001) of parallel observations at the station of Rovereto to estimate the potential advantage of using cloud cover in a homogenization algorithm (a thorough evaluation of the homogenization performance requires multiple independent parallel records and is beyond the scope of this paper).

In this record, traditional liquid-in-glass instruments in a "meteorological window" (wooden wall screen with NNW exposure, 7.5 m from the ground) are compared to the electronic sensor of an automatic weather station (AWS). The sensor was installed in a modern, freestanding screen over grassy terrain (Figure 2).

There is a discrepancy between the observation times of daily extremes at the meteorological window (T_n of the previous 24 h observed at 14:00 LT, T_x observed at 19:00 LT) and at the AWS (calendar day). This change in the definition of the meteorological day is commonly found in long records (see section 3.2).

Meteorological windows (i.e., instruments installed outside a window on a north facing wall, usually contained in wooden or metallic radiation screens of various kinds), were employed by the Austrian, Italian, and Swiss weather services in the late nineteenth century [*Jelinek*, 1869; *Denza*, 1882; *Parker*, 1994; *Begert et al.*, 2005; *Brunetti et al.*, 2006]. They remained a very common configuration for meteorological observatories well into

Table 2. Parameters Used for the Evaluation of the Homogenization Performance

Parameter	Definition	Unit
$\bar{E}$	Mean error	°C
E_{RMS}	Root-mean-square error	°C
E_5	5th percentile of the error distribution	°C
E_{95}	95th percentile of the error distribution	°C
$\%E_{0.5}$	Fraction of absolute errors $\leq$ 0.5°C	%
E_M	Error for the absolute maximum value	°C
E_m	Error for the absolute minimum value	°C
$n_{E<5}$	Absolute error in the number of values lower than the 5th percentile	Days
$n_{E>95}$	Absolute error in the number of values higher than the 95th percentile	Days

the twentieth century, especially in Italy. Three of the stations that we analyze were initially equipped with a meteorological window. In Domodossola, the transition to a freestanding wooden screen took place in 1905, while metadata suggest that a meteorological window might have been used as late as 1980 in Riva. Finally, Rovereto is an exceptional case, as a meteorological window is still in operation here (see section 2).

We simulate an inhomogeneity in the record of Rovereto by switching from the meteorological window series to the AWS series on 1 January 1997. Therefore, we use the period 1992–1996 as h_1 and 1997–2001 as h_2, with h_1 corrected and then compared to the parallel record of the AWS for validation. We correct the inhomogeneity using our homogenization algorithm with and without the aid of cloud cover (denoted by CLOUD and NO-CLOUD, respectively). Note that both CLOUD and NO-CLOUD apply the same corrections for C_2. We use Lugano, Riva, Locarno-Monti, Mezzolombardo, and Trento-Laste as reference series (the record of Domodossola does not have enough data in the period h_2).

Moreover, we apply two additional state-of-the-art homogenization methods, specifically developed for daily temperature data: the higher-order moments method (HOMAD) from *Toreti et al.* [2012] and the percentile-matching algorithm (PM95) from *Trewin* [2013]. The former method uses only one reference series (Riva).

To evaluate the performance of the homogenization in this case study, we define various parameters related to errors in the whole sample and in the extremes. The parameters are described in Table 2. The percentiles used for the definitions of the parameters are calculated from the empirical distribution of the observations made at the AWS during 1992–2001.

3.4. Definition of Temperature Extremes and Trends

In order to allow for a direct comparison, we adopt the same definitions used in *Della-Marta et al.* [2007a] for the duration and intensity of heat waves. Specifically, we define a hot day as a day where T_x exceeds its long-term daily 95th percentile. The daily percentiles are calculated from the empirical distribution of the values observed in a 15 day window around the target day, in the period 1901–2000. The hot day index (HD) is the number of hot days within a June–August season expressed as a percentage of time, while a heat wave (HW) is the maximum number of consecutive hot days that occur in the season.

Similarly, we define a cold day as a day where T_n is below its long-term daily 5th percentile, the cold days index (CD) as the number of cold days within a December–February season expressed as a percentage of time and a cold spell (CS) as the maximum number of consecutive cold days in the season.

These indices do not follow a normal distribution and usually have outliers. To calculate their trends, we employ a robust linear regression using a so-called *M*-estimator [*Venables and Ripley*, 2002], specifically by minimizing the Huber's loss function [*Huber and Ronchetti*, 2009] instead of the squared residuals as in ordinary least squares. We verified that the results are not significantly different from those obtained using the median of pairwise slopes [*Sen*, 1968], another robust regression technique commonly applied in climate research.

To evaluate the statistical significance of trends, we use the Mann-Kendall nonparametric test [*Sneyers*, 1990].

Table 3. Performance of the Homogenization Methods for the Parallel Record of Rovereto

		$\bar{E}$		E_{RMS}		E_5		E_{95}		$\%E_{0.5}$		E_M		E_m		$n_{E<5}$		$n_{E>95}$	
Method	Reference[a]	T_x	T_n	T_x	T_n	T_x	T_n	T_x	T_n	T_x	T_n	T_x	T_n	T_x	T_n	T_x	T_n	T_x	T_n
Winter[b]																			
Uncorrected	—	−1.98	1.82	2.56	2.04	−4.6	0.5	0.4	3.3	21.9	4.9	−2.6	1.0	−0.7	1.8	20	−15	−8	32
HOMAD	Riva	−0.55	0.34	1.69	1.08	−3.0	−1.3	1.9	2.0	19.0	54.0	1.3	0.0	0.6	0.0	0	−5	−6	−1
PM95	All	0.25	0.49	1.65	1.06	−2.2	−0.8	2.8	2.1	26.1	52.0	1.9	0.0	0.0	−0.6	−1	−3	−2	3
NO-CLOUD	All	0.37	0.25	1.64	0.96	−2.2	−1.1	2.8	1.8	22.6	57.7	0.0	−0.4	1.4	0.4	−14	−5	−6	−5
CLOUD	All	0.29	0.35	1.15	0.97	−1.6	−0.7	1.9	2.1	31.4	63.7	0.7	0.1	1.4	0.3	−6	−3	−5	3
CLOUD	Mezzolombardo	−0.22	0.34	1.14	0.97	−2.0	−0.9	1.4	1.8	42.3	55.5	0.4	−0.3	0.7	0.3	0	−4	−6	−4
CLOUD	Riva	0.16	0.11	1.14	0.91	−1.8	−0.9	1.7	1.9	32.7	68.1	0.4	−0.2	1.2	−0.1	−5	3	−6	0
CLOUD	Trento-Laste	1.09	0.69	1.57	1.16	−0.7	−0.5	2.8	2.4	19.2	36.5	1.7	0.3	1.9	0.8	−13	−6	4	8
CLOUD	Locarno-Monti	0.16	0.32	1.15	0.97	−1.8	−0.7	1.7	2.3	30.5	65.5	0.5	0.2	1.5	0.2	−6	1	−6	7
CLOUD	Lugano	0.21	0.28	1.15	0.96	−1.7	−0.7	1.8	2.2	31.6	69.5	0.7	0.3	1.4	0.0	−5	0	−5	6
CLOUD+	Lugano	0.33	0.25	1.24	0.98	−1.5	−0.8	2.1	2.1	35.6	67.0	1.3	0.3	1.0	0.0	−3	1	−3	7
Summer																			
Uncorrected	—	−1.55	2.21	1.89	2.35	−3.3	1.1	0.2	3.4	15.2	0.4	−1.6	1.5	0.5	3.6	6	−20	−25	101
HOMAD	Riva	−0.14	0.32	1.18	0.96	−2.0	−0.9	1.8	1.7	38.3	50.9	−0.1	−0.7	2.0	2.0	−7	−9	−17	−3
PM95	All	−0.18	0.19	1.06	0.84	−1.9	−0.9	1.5	1.4	38.9	59.3	0.7	−0.2	1.9	1.3	−5	7	−5	3
NO-CLOUD	All	0.18	0.10	0.94	0.82	−1.3	−0.9	1.7	1.3	41.5	56.1	0.6	−0.6	1.7	1.4	−4	6	−2	1
CLOUD	All	0.16	0.06	0.93	0.81	−1.4	−0.9	1.6	1.2	38.9	57.4	0.8	−0.7	1.8	1.3	−3	6	−2	0
CLOUD	Mezzolombardo	0.81	0.42	1.24	0.92	−0.7	−0.7	2.3	1.6	28.0	48.3	1.5	−0.3	2.4	2.0	−10	2	10	16
CLOUD	Riva	−0.26	0.10	0.97	0.81	−1.8	−1.0	1.1	1.3	39.1	57.4	0.2	−0.6	1.6	1.4	−1	7	−10	0
CLOUD	Trento-Laste	−0.10	−0.43	0.92	0.94	−1.6	−1.4	1.3	0.7	38.9	40.9	0.6	−0.6	1.2	0.9	−1	21	−3	−7
CLOUD	Locarno-Monti	0.19	0.08	0.97	0.83	−1.4	−1.0	1.6	1.3	37.0	57.6	0.8	−0.6	2.0	1.0	−6	6	−2	−1
CLOUD	Lugano	0.15	0.18	0.95	0.85	−1.5	−0.9	1.6	1.4	37.6	57.2	0.9	−0.5	1.8	1.0	−4	3	−2	2
CLOUD+	Lugano	−0.01	0.15	1.00	0.91	−1.6	−0.9	1.5	1.5	35.2	57.8	1.0	−0.1	0.8	1.1	0	1	−1	4

[a]Reference series that have been used for each method.
[b]Winter (summer) is defined here as the months from December to February (June to August).

4. Results

4.1. Homogenization Performance

Table 3 summarizes the results of the homogenization for each of the methods that we tested, for winter and summer. The mean errors ($\bar{E}$) in the uncorrected data are consistent with other studies [e.g., *Parker*, 1994; *Böhm et al.*, 2010] in that *DTR* is lower for the meteorological window than for a freestanding screen; T_x (T_n) is, over the whole year, on average 2°C lower (higher) in the meteorological window.

The largest impact of the use of cloud cover is found for T_x in winter, when E_{RMS} for CLOUD is 30% lower than for NO-CLOUD. For T_n, there are no improvements in E_{RMS}, nevertheless $\%E_{0.5}$ is higher for CLOUD. HOMAD and PM95 produce a higher E_{RMS} than both CLOUD and NO-CLOUD for all combinations of variables and seasons.

The absolute errors in the counts of the extremes ($n_{E<5}$ and $n_{E>95}$) are, on average, 43% lower in winter and 15% lower in summer when using CLOUD instead of NO-CLOUD. HOMAD and PM95 have smaller errors in the counts than CLOUD in winter (−29% and −47%, respectively) despite their higher E_{RMS} but larger errors in summer (+227% and +82%).

To evaluate the influence of the number of reference series, we also show the results obtained by using each one of the reference series alone. The performances hardly depend on the reference series, with some exceptions; Trento-Laste and Mezzolombardo provide significantly worse corrections than the other reference series, indicating that their records have small inhomogeneities between 1992 and 2001 that were not detected by the homogenization algorithm. Nevertheless, they still improve the homogeneity of the candidate series and provide better results than NO-CLOUD, HOMAD, and PM95 for some of the indices.

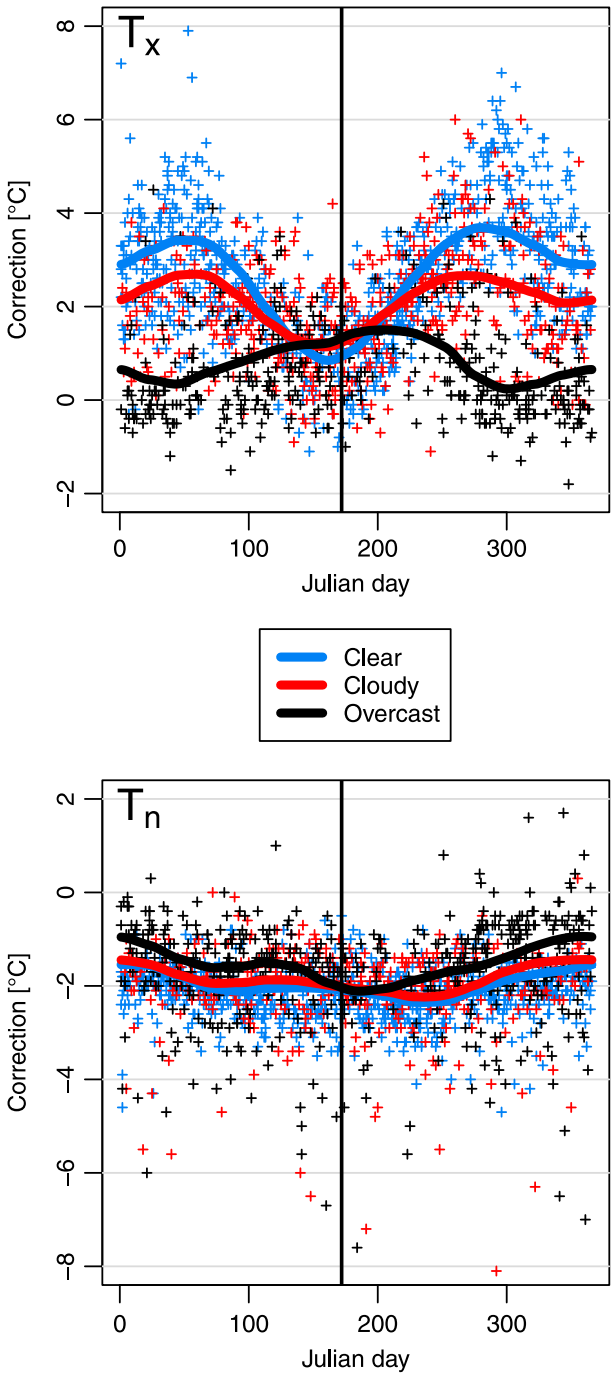

Figure 5. Estimated corrections (lines) using the CLOUD method for the series of Rovereto over the period 1992–1996 as a function of the Julian day for T_x (top) and T_n (bottom). Colors represent the three cloud cover categories; points represent observed differences in the parallel record. The vertical lines locate the summer solstice.

For Lugano we also tested a variation of CLOUD, which we call CLOUD+, where the corrections are calculated using only days when a certain cloud cover category is observed at both the candidate and the reference station. As expected, the reduction of the sample size counteracts the improved similarity between cloud cover at the two stations, resulting in better performances for some indices, but worse performances for others.

The results from the single reference series allow for a better comparison with HOMAD, which uses only the series of Riva, and confirm the better performances of CLOUD for most of the indices.

Table 4. Linear Trends of Annual Mean Temperatures Over the Period 1874–2015[a]

	T_x		T_n		*DTR*		T_m	
Station	Before	After	Before	After	Before	After	Before	After
Domodossola	0.9 ± 0.2	1.0 ± 0.2	-0.1 ± 0.2	1.7 ± 0.1	0.6 ± 0.3	-0.9 ± 0.1	0.6 ± 0.2	1.4 ± 0.1
Lugano	-0.1 ± 0.2	1.5 ± 0.1	2.2 ± 0.1	1.7 ± 0.1	-2.3 ± 0.2	-0.2 ± 0.1	1.0 ± 0.1	1.6 ± 0.1
Riva	1.0 ± 0.3	0.6 ± 0.1	-1.5 ± 0.3	1.8 ± 0.1	2.6 ± 0.2	-1.2 ± 0.1	-0.2 ± 0.3	1.2 ± 0.1
Rovereto	1.6 ± 0.2	1.2 ± 0.1	1.7 ± 0.1	1.3 ± 0.1	-0.2 ± 0.2	-0.1 ± 0.1	1.7 ± 0.1	1.3 ± 0.1
Southern Alps[b]	0.7 ± 0.2	1.1 ± 0.1	0.7 ± 0.2	1.6 ± 0.1	-0.1 ± 0.2	-0.5 ± 0.1	0.7 ± 0.1	1.4 ± 0.1

[a]In terms of least squares regression coefficients (°C century^{-1}), with standard errors, before and after the homogenization.
[b]Calculated from the average of the anomaly series of the four stations.

Figure 5 shows the daily corrections applied to each C_k in CLOUD. For T_n, the influence of cloud cover on the corrections is rather small. On the other hand, T_x corrections have a strong annual cycle for C_1 and C_2, resembling the annual cycle of the solar elevation angle, while corrections for C_3 are close to zero for most of the year. The inverse correlation between solar elevation and correction can be explained by a larger fraction of shortwave radiation reaching the meteorological window when the Sun is higher (mostly due to reflection by the surrounding environment, but even through direct radiation in the late-afternoon hours between June and July). This causes an overestimation of T_x with respect to winter months because of the poor radiation-shielding characteristics of the wooden case.

The corrections for C_3 are very close to those for the other categories in late spring and summer, probably as a consequence of the larger variability in space and time of cloud cover rather than a real physical signal. Moreover, corrections for C_3 are more difficult to estimate in summer because overcast conditions are much less frequent, particularly in July and August.

Figure 5 also shows the "real" corrections observed for each day in the period 1992–1996 (i.e., temperature at the weather station minus temperature at the meteorological window). The differences for C_3 (black points) have a larger variability in spring and summer than for the other categories, indicating that they are often contaminated by rapid changes in cloud cover, which are typical of those seasons (see Figure S6). A number of outliers are also evident, particularly for T_n. They are mostly caused by the different definitions of the meteorological day (see section 3.3.3); when there is a disruption of the normal daily cycle of temperature (due to, e.g., a thunderstorm), a daily extreme can be assigned to different days depending on the definition adopted and can sometime be assigned to 2 days in a row when it is reached near the end of a day (e.g., near midnight for the AWS). The homogenization procedure cannot correct these inconsistencies but can correct the mean bias that they cause (as discussed in section 3.2).

4.2. Trends in Annual Means

Table 4 summarizes the impact of the homogenization on the trends of T_x, T_n, *DTR*, and mean temperature ($T_m = \frac{T_x+T_n}{2}$) over the entire analyzed period. Trends are more similar between stations after the homogenization, although they are not directly comparable because of the data gaps in the records (only years with at least 350 available daily values are used for trend calculation). The mean regional trend of T_n is nearly 50% larger than that of T_x, causing a significant negative trend for *DTR*. All homogenized trends in Table 4 are statistically significant at the 1% level ($p < 0.01$), except the trends of *DTR* in Lugano ($p < 0.05$) and Rovereto ($p < 0.20$).

Figures 6 and 7 show the annual mean T_x and T_n before and after the homogenization for each of the analyzed records. The corrections can be very high, up to 4°C, as in the case of T_n in Domodossola in the early years (for clear days they can reach 6.5°C in winter, not shown). Such large corrections are justified when meteorological windows are employed because of the height from the ground (up to 22 m in Domodossola) and the heat dispersion of the building in which the window is installed.

Figures 6 and 7 (fifth panel) show the average regional series (calculated from the anomaly series of the four stations), which emphasizes a larger decadal variability in T_x than in T_n. This contributes to a reduced trend in T_x because of the particular time frame analyzed, which begins with a warm phase, but does not, on its own, manage to explain the large difference with the trend in T_n.

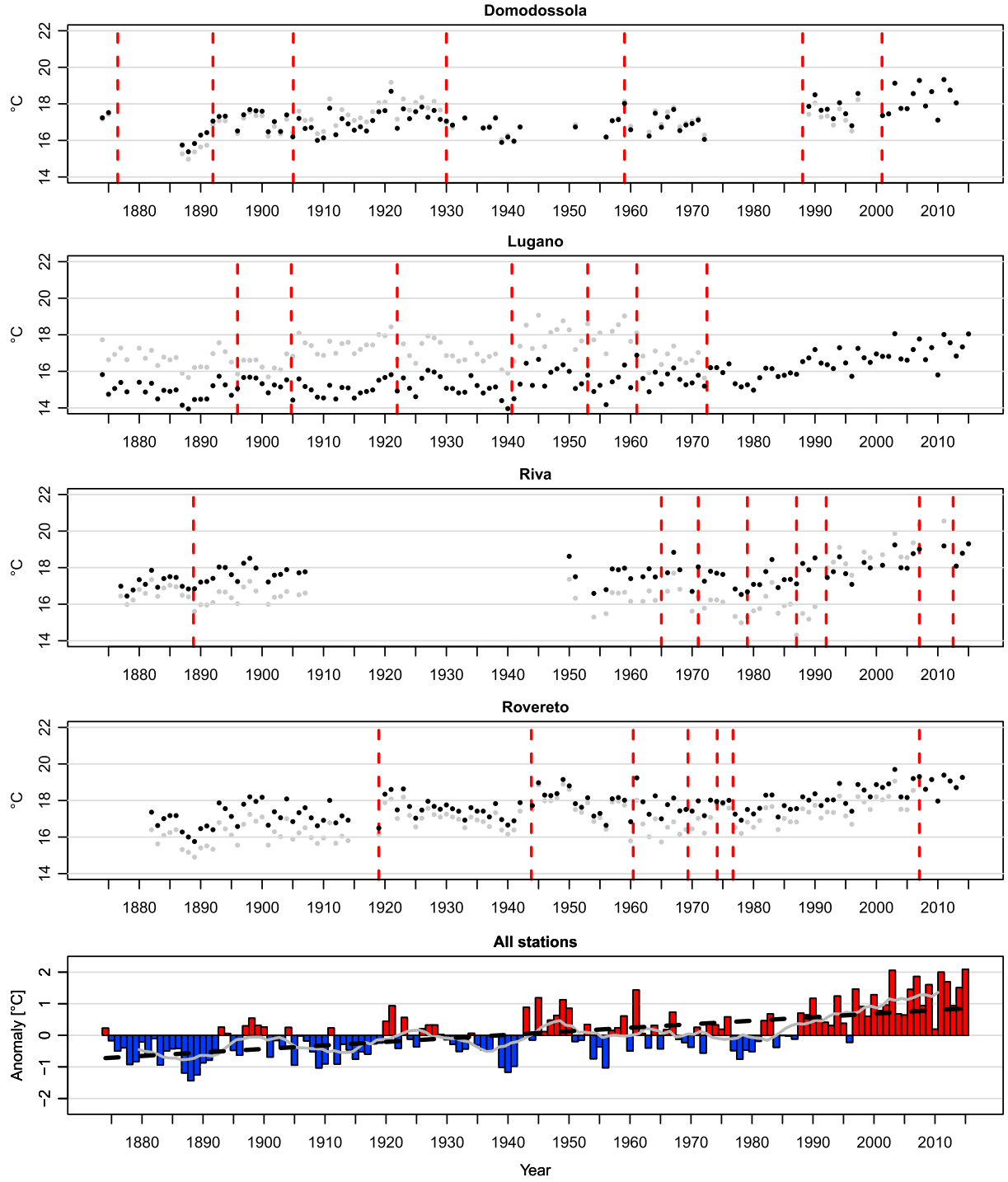

Figure 6. Annual mean T_x series before (grey dots) and after (black dots) homogenization (only years with at least 350 available daily observations are plotted). Vertical dashed lines indicate the breakpoints. (fifth panel) The average of the four homogenized series (anomalies with respect to 1901–2000), together with the 11 year moving average (grey solid line) and the linear trend (dashed black line).

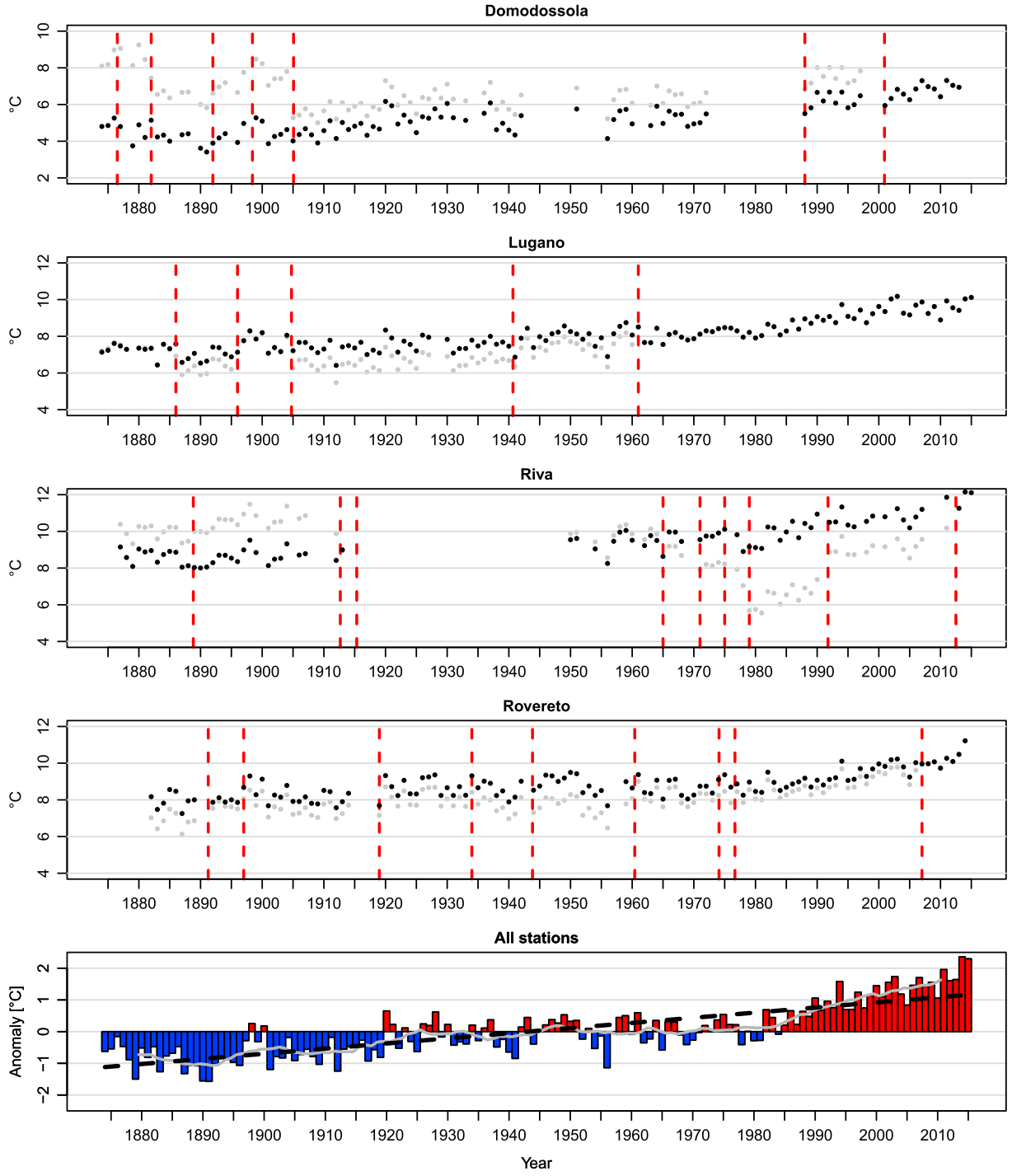

Figure 7. Same as Figure 6 for T_n.

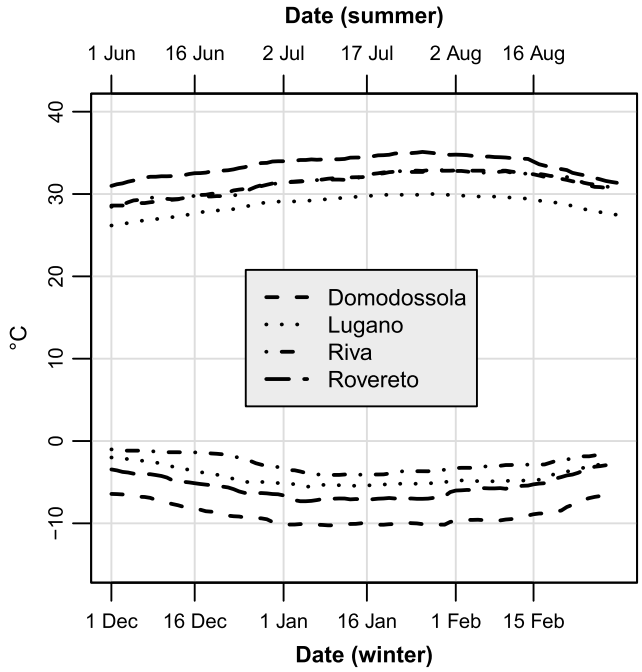

Figure 8. Daily thresholds used for the calculation of the extreme indices: 95th percentiles of T_x in summer (upper lines) and 5th percentiles of T_n in winter (lower lines).

4.3. Trends in Extreme Indices

Figure 8 shows the thresholds used for the calculation of the extreme indices (see section 3.4). The mitigation effects of Lake Lugano and Lake Garda are evident in the thresholds of Lugano and Riva, whereas the stations of Domodossola and Rovereto have a more continental climate (i.e., larger differences between cold and warm extremes).

4.3.1. Heat Waves

Results for the indices HD and HW are summarized in Figures 9 and 10. There are some minor discrepancies between the four records; on the one hand, these can be real physical signals caused by the different microclimates, but on the other hand, it is likely that a few significant errors still affect the data. For example, a few spikes for HD in the record of Lugano (most notably those of 1881 and 1904) are not seen in the other records. Nevertheless, both HD and HW are characterized by positive trends for all stations, statistically significant at the 1% level. Similarly to the mean values, we calculated a regional series of the extreme indices from the four anomaly series. The resulting trends are 6.8 ± 0.9 days century^{-1} for HD and 2.7 ± 0.4 days century^{-1} for HW.

The year 2003 has by far the highest value of HD, with HW exceeding 12 days for all stations. The recent sequence of heat waves that characterized the summer of 2015 in Europe brought the second highest value of HD since 1874 at the stations of Lugano and Riva (data for 2015 were not available at the other stations). The 2015 value of HW matched the 2003 value in Lugano (13 days). A summer without hot days has not occurred for almost 30 years, whereas it is found on average every third year before 1980.

4.3.2. Cold Spells

Results for the indices CD and CS are summarized in Figures 11 and 12. All stations show similar and highly significant ($p < 0.01$) negative trends. In this case, the most significant discrepancy comes from the record of Riva in the 1980s. The regional mean trends are -4.0 ± 0.7 days century^{-1} for CD and -2.0 ± 0.4 days century^{-1} for CS.

After a step-like reduction in the early 1970s, cold days have become rare in the last few decades. Since 1980, only the winter of 1984/1985 had a remarkable cold spell (CS > 5 at all available stations). The exceptionally cold winter of 1928/1929 had more cold days than the whole period 1975–2015 in Lugano or 1986–2014 in Rovereto. On average, we observe a reduction of about 80% for CD and 70% for CS between the periods 1881–1910 and 1981–2010.

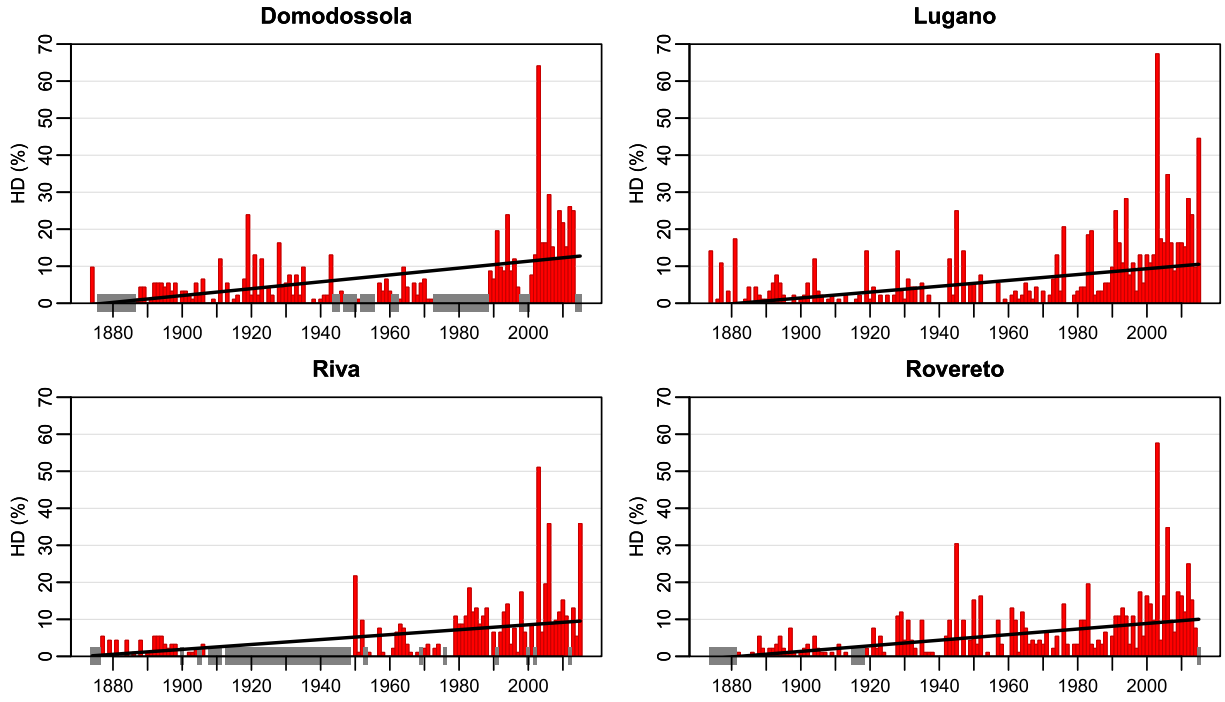

Figure 9. Time series for HD. Missing years are marked by the grey shading on the time axis (at least 95% of nonmissing daily data are required to calculate the index for a certain year). Lines represent linear trends.

5. Discussion

The parallel record of Rovereto supports our hypothesis that cloud cover can be used to improve the quality of a statistics-based homogenization of daily temperature records. In the specific case of Rovereto, shortwave radiation (direct and scattered) clearly plays a large role in shaping daily differences between the two configurations, while the impact of longwave radiation is much smaller. We expect a strong variability of the homogenization performance with other kinds of inhomogeneities and in different climates; hence, very different behaviors are possible for other breakpoints and other stations. For example, the corrections used to homogenize our data set also show a large influence of cloud cover for T_n for some inhomogeneities (not shown), and we can then expect a significant improvement of the homogenization performance for T_n in those cases. Obviously, other inhomogeneities are not influenced by cloud cover (e.g., the recalibration of a thermometer), and the use of our method will not bring any improvement with respect to other methods. In a worst-case scenario, represented by very short h_1 and/or h_2 subperiods, it could even produce a worse homogenization performance due to the reduction in sample size. Small sample sizes, related to the limited number of reference series available in this study, also prevent the correction of high-order moments of the distribution within a cloud cover category, which could improve the performance of the homogenization when cloud cover variability is high (e.g., in summer) and when other variables (e.g., wind) play an important role.

The regional trend that we obtained for the mean temperature is compatible with previous analyses for the southern part of the Greater Alpine Region [*Auer et al.*, 2007; *Brunetti et al.*, 2009]. The magnitudes of trends in the original data were on average underestimated for all variables, although opposite behaviors are found for specific stations depending on the type of radiation screen used in the first part of the record. In particular, the transition from the meteorological window to modern screens causes a strong underestimation of the trends for T_n and, consequently, an overestimation of the trends for *DTR*.

While a negative trend in *DTR* has been observed in most parts of the world during the second half of the twentieth century [e.g., *Vose et al.*, 2005], studies over longer periods are rare. *Moberg et al.* [2006] found no difference between the trends in T_x and T_n over the period 1901–2000 in Europe as a whole. A stronger trend for T_n was observed by *Brunetti et al.* [2006] for the southern Alps over the period 1865–2003; they did not find a similar feature in the adjoining Po Plain. In the northern Alps, a significant negative trend for *DTR* was found for low-altitude stations over the period 1901–1990, but not for mountain stations [*Weber et al.*, 1997].

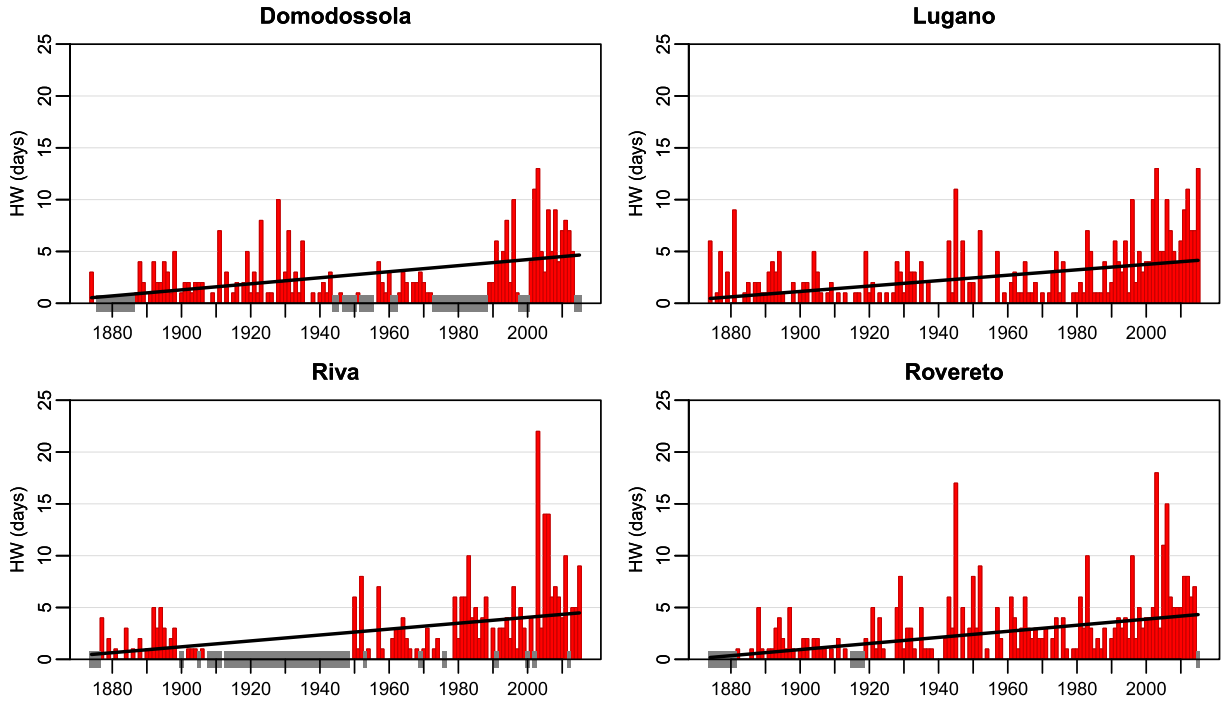

Figure 10. Same as Figure 9 for HW.

On the other hand, *Acquaotta et al.* [2014] found stronger positive trends for T_n than for T_x at high-altitude stations in the western Italian Alps over the period 1961–2010. Note that we use only low-altitude stations in the present study, as did *Brunetti et al.* [2006] for their Alpine subset.

DTR is negatively correlated with cloud cover [e.g., *Dai et al.*, 1999]; therefore, a positive trend in cloud cover would explain at least part of the trend in *DTR*. Previous studies [*Auer et al.*, 2007; *Brunetti et al.*, 2009] did find

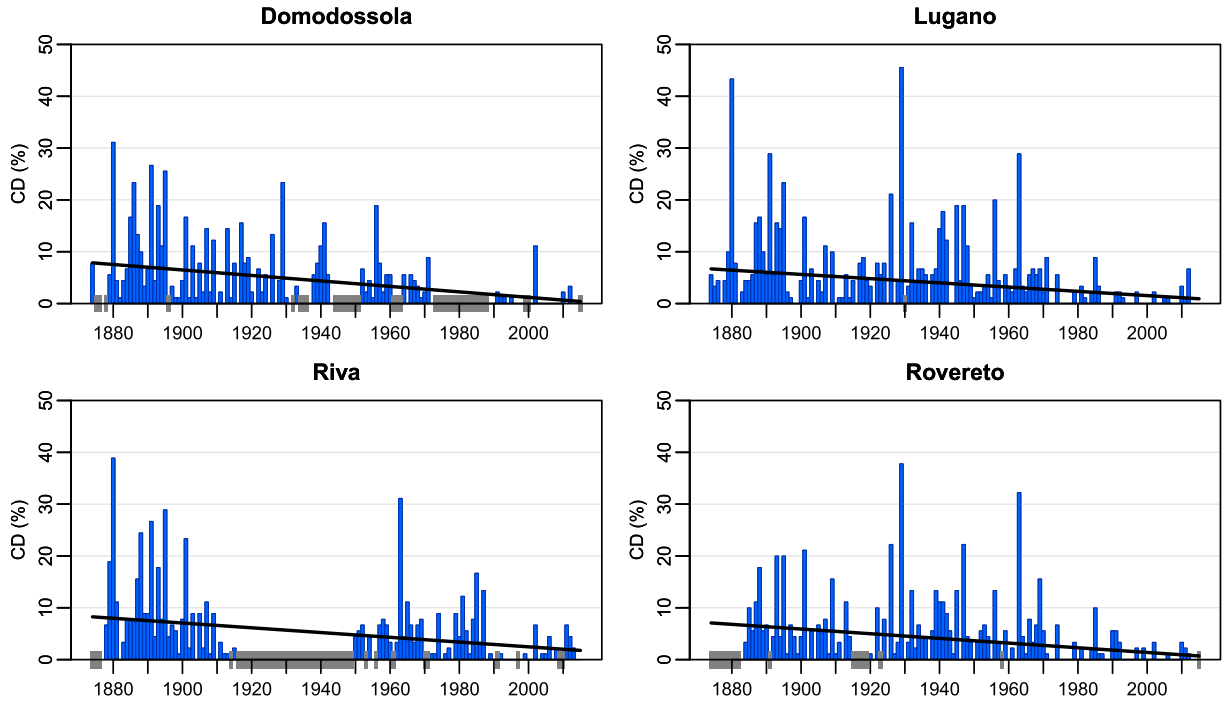

Figure 11. Same as Figure 9 for CD.

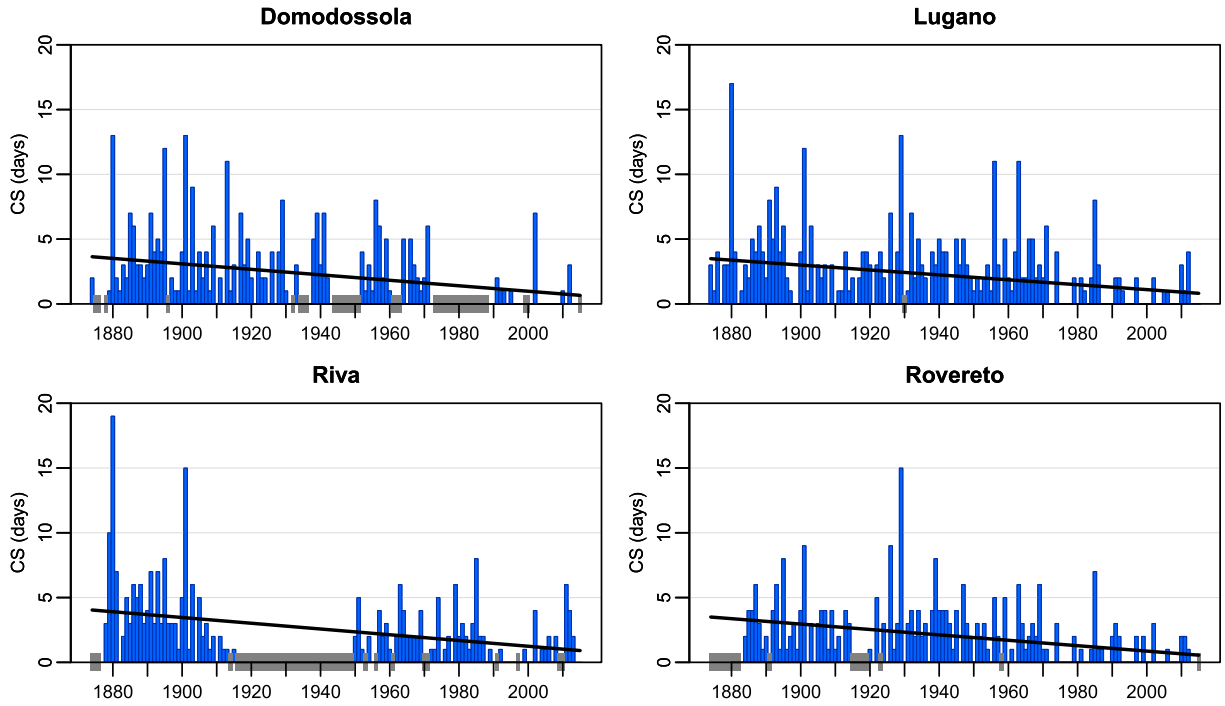

Figure 12. Same as Figure 9 for CS.

an increase of cloud cover in the southern Alps, but with a low statistical significance in all seasons except summer. The only complete cloud cover record in our data set, Lugano, shows a strong positive trend (1.0 ± 0.1 tenths century^{-1}), while that of Rovereto (1919–2013) shows the opposite (-0.9 ± 0.1 tenths century^{-1}). The homogeneity of these records is a big concern, and no solid conclusions can be drawn.

Our results for the warm extreme indices are consistent with those of *Della-Marta et al.* [2007a], who found positive trends between 1 and 9 days century^{-1} for HD and between 0.3 and 3 days century^{-1} for HW in the Alps. This means that over the last ca. 140 years, HD and HW have increased on average by more than 200%. However, if we restrict the analysis to the period 1880–2005 (i.e., the period analyzed in *Della-Marta et al.* [2007a]), then we find significantly smaller trends than for our complete period. As evident from Figures 9 and 10, the increase has taken place mostly over the last 20 to 40 years, consistent with the changes in mean T_x (Figure 6). This rapid increase in the trends is probably related in part to natural variability [e.g., *Sutton and Hodson*, 2005; *Della-Marta et al.*, 2007b] that enhanced the anthropogenic warming [*Stocker et al.*, 2014]. Soil moisture has also been shown to have an important influence on heat waves [e.g., *Mueller and Seneviratne*, 2012], and the significant decrease of winter and spring precipitation observed in the southern Alps over the last few decades [*Brugnara et al.*, 2012] might be an additional contributor to the recent trends observed in the warm extreme indices.

A strong reduction of winter CD after ca. 1970 was also found by *Yan et al.* [2002] for two stations in the Po Plain (Milan and Padua). However, they found much lower values than those in our stations over the period 1900–1930, which caused the secular trends to be nonsignificant. A similar behavior was observed in south-eastern Europe and was linked in part to changes in the frequency and residence time of synoptic circulation patterns [*Domonkos et al.*, 2003]. In fact, low T_n values at low elevations are usually generated by a strong temperature inversion, which can only occur with stable anticyclonic conditions that persist for several days.

6. Conclusions

In this work, we responded to the need for additional meteorological records at daily resolution in climate change research [e.g., *Menne et al.*, 2012; *Donat et al.*, 2013], by presenting three maximum and minimum temperature series for the southern Alps dating back to the second half of the nineteenth century. We also recovered subdaily cloud cover observations for the same stations and used them as a constraint for

the homogenization of the temperature series, in order to take into account the effect of radiation on the magnitude of the inhomogeneities on the daily scale.

The homogenization algorithm that we proposed requires cloud cover observations only for the candidate station, under the assumption that spatial variability is sufficiently low. Future improvements in the method may include the use of wind, snow cover, or other meteorological parameters, as well as the correction of high-order moments of the distribution within a certain cloud cover category.

A parallel record at one of the stations allowed for a validation of the homogenization procedure for a specific inhomogeneity (transition from a meteorological window to a modern weather station). The results indicate that the use of cloud cover significantly improves the homogenization performance, although a more solid validation based on multiple parallel records and/or physics-based models is needed, in particular, for climatic regions where cloud cover variability is larger than in our study area (e.g., northern Europe).

We obtained long-term trends in the mean annual temperature that are consistent with previous studies. This includes our finding of a significant negative trend in the daily temperature range, probably related in part to an increase in cloud cover.

We defined four extreme indices to represent frequency and duration of heat waves in summer and cold spells in winter. Warm extremes have strongly increased, particularly over the last three decades, while a reduction of similar magnitude has affected cold extremes.

Our results clearly indicate that the southern Alps have undergone an important change toward a warmer climate that has also affected extreme temperatures. However, the homogenization of daily temperature observations poses greater difficulties than the homogenization of monthly or annual means, and therefore, the trends in extreme indices are affected by a larger uncertainty. The Alpine region is particularly rich in long meteorological records, many of them not yet digitized, meaning that a considerable potential still exists to reduce this uncertainty.

Acknowledgments
This work was supported by the EU Horizon 2020 EUSTACE project (grant agreement 640171). Renate Auchmann also acknowledges the SNF project TWIST (200021_146599/1). Parts of the records of Riva and Rovereto were digitized from material recovered by the project ASTRO (http://cma.entecra.it/astro) or were downloaded from the website of the meteorological service of the Autonomous Province of Trento (Meteotrentino: http://www.meteotrentino.it) as were the series of Mezzolombardo and Trento-Laste. The record of Lugano and the other Swiss series were provided by the Swiss National Weather Service (Federal Office of Meteorology and Climatology MeteoSwiss) through the IDAweb portal (http://gate.meteoswiss.ch/idaweb). ECA and D series were downloaded from KNMI's website (http://eca.knmi.nl). We thank Andrea Toreti for providing the HOMAD software, Michael Begert for providing additional metadata for the station of Lugano, and Blair Trewin for fruitful discussions. We also thank the three anonymous reviewers, as well as the Associate Editor, for their feedbacks. The preservation and digitization of the record of Rovereto is made possible thanks to the support of the Fondazione Museo Civico di Rovereto. Alessio Bozzo wishes to thank and remember Fausto Maroni who was responsible for the maintenance of the meteorological observatory in Rovereto. He took care of the daily weather observations over the past 15 years and digitized the daily records used in this work. The weather observations in the historical site can continue today, thanks to his dedication. A thank you also goes to Filippo Orlando, who succeeded as director of the observatory.

References

Acquaotta, F., S. Fratianni, and D. Garzena (2014), Temperature changes in the North-Western Italian Alps from 1961 to 2010, *Theor. Appl. Climatol.*, *122*, 619–634, doi:10.1007/s00704-014-1316-7.

Auchmann, R., and S. Brönnimann (2012), A physics-based correction model for homogenizing sub-daily temperature series, *J. Geophys. Res.*, *117*, D17119, doi:10.1029/2012JD018067.

Auer, I., et al. (2007), HISTALP—Historical instrumental climatological surface time series of the Greater Alpine Region, *Int. J. Climatol.*, *27*(1), 17–46, doi:10.1002/joc.1377.

Begert, M., T. Schlegel, and W. Kirchhofer (2005), Homogeneous temperature and precipitation series of Switzerland from 1864 to 2000, *Int. J. Climatol.*, *25*(1), 65–80, doi:10.1002/joc.1118.

Böhm, R., P. D. Jones, J. Hiebl, D. Frank, M. Brunetti, and M. Maugeri (2010), The early instrumental warm-bias: A solution for long central European temperature series 1760–2007, *Clim. Change*, *101*(1–2), 41–67, doi:10.1007/s10584-009-9649-4.

Brugnara, Y., M. Brunetti, M. Maugeri, T. Nanni, and C. Simolo (2012), High-resolution analysis of daily precipitation trends in the central Alps over the last century, *Int. J. Climatol.*, *32*(9), 1406–1422, doi:10.1002/joc.2363.

Brunet, M., O. Saladié, P. Jones, J. Sigró, E. Aguilar, A. Moberg, D. Lister, A. Walther, and C. Almarza (2008), A case-study/guidance on the development of long-term daily adjusted temperature datasets, *WMO-TD-1425*, World Meteorological Organization, Geneva, Switzerland.

Brunetti, M., M. Maugeri, F. Monti, and T. Nanni (2006), Temperature and precipitation variability in Italy in the last two centuries from homogenised instrumental time series, *Int. J. Climatol.*, *26*(3), 345–381, doi:10.1002/joc.1251.

Brunetti, M., G. Lentini, M. Maugeri, T. Nanni, I. Auer, R. Böhm, and W. Schöner (2009), Climate variability and change in the Greater Alpine Region over the last two centuries based on multi-variable analysis, *Int. J. Climatol.*, *29*(15), 2197–2225, doi:10.1002/joc.1857.

Caussinus, H., and O. Mestre (2004), Detection and correction of artificial shifts in climate series, *J. R. Stat. Soc.*, *53*(3), 405–425, doi:10.1111/j.1467-9876.2004.05155.x.

Cleveland, W. S., and S. J. Devlin (1988), Locally weighted regression: An approach to regression analysis by local fitting, *J. Am. Stat. Assoc.*, *83*(403), 596–610, doi:10.1080/01621459.1988.10478639.

Cocheo, C., and D. Camuffo (2002), Corrections of systematic errors and data homogenisation in the daily temperature Padova series (1725–1998), *Clim. Change*, *53*, 77–100, doi:10.1023/A:1014950306015.

Dai, A., K. E. Trenberth, and T. R. Karl (1999), Effects of clouds, soil moisture, precipitation, and water vapor on diurnal temperature range, *J. Clim.*, *12*(8), 2451–2473, doi:10.1175/1520-0442(1999)012<2451:EOCSMP>2.0.CO;2.

Della-Marta, P., and H. Wanner (2006), A method of homogenizing the extremes and mean of daily temperature measurements, *J. Clim.*, *19*(17), 4179–4197, doi:10.1175/JCLI3855.1.

Della-Marta, P. M., M. R. Haylock, J. Luterbacher, and H. Wanner (2007a), Doubled length of western european summer heat waves since 1880, *J. Geophys. Res.*, *112*, D15103, doi:10.1029/2007JD008510.

Della-Marta, P. M., J. Luterbacher, H. von Weissenfluh, E. Xoplaki, M. Brunet, and H. Wanner (2007b), Summer heat waves over western Europe 1880–2003, their relationship to large-scale forcings and predictability, *Clim. Dynam.*, *29*(2–3), 251–275, doi:10.1007/s00382-007-0233-1.

Denza, P. F. (1882), *Istruzioni per le Osservazioni Meteorologiche e Per L'Altimetria Barometrica*, Collegio Artigianelli Tip., Turin, Italy.

Di Napoli, G., and L. Mercalli (2008), *Il Clima di Torino: Tre Secoli di Osservazioni Meteorologiche*, 936 pp., SMS, Turin, Italy.

Domonkos, P., J. Kysely, K. Piotrowicz, P. Petrovic, and T. Likso (2003), Variability of extreme temperature events in south-central europe during the 20th century and its relationship with large-scale circulation, *Int. J. Climatol.*, *23*(9), 987–1010, doi:10.1002/joc.929.
Donat, M., et al. (2013), Updated analyses of temperature and precipitation extreme indices since the beginning of the twentieth century: The HadEX2 dataset, *J. Geophys. Res.*, *118*, 2098–2118, doi:10.1002/jgrd.50150.
Edwards, P. N. (2004), "A vast machine": Standards as social technology, *Science*, *304*(5672), 827–828, doi:10.1126/science.1099290.
European Environment Agency (2009), Regional climate change and adaptation: The Alps facing the challenge of changing water resources,
EEA Rep. No 8/2009, Eur. Environ. Agency, Copenhagen, doi:10.2800/12552.
Frei, C., and C. Schär (1998), A precipitation climatology of the Alps from high-resolution rain-gauge observations, *Int. J. Climatol.*, *18*(8), 873–900, doi:10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9.
Gardner, A. S., et al. (2013), A reconciled estimate of glacier contributions to sea level rise: 2003 to 2009, *Science*, *340*(6134), 852–857, doi:10.1126/science.1234532.
Giovannini, L., L. Laiti, D. Zardi, and M. de Franceschi (2015), Climatological characteristics of the Ora del Garda wind in the Alps, *Int. J. Climatol.*, *35*(14), 4103–4115, doi:10.1002/joc.4270.
Huber, P. J., and E. M. Ronchetti (2009), *Robust Statistics*, 2nd ed., Wiley, New York.
Huwald, H., C. Higgins, M.-O. Boldi, E. Bou-Zeid, M. Lehning, and M. Parlange (2009), Albedo effect on radiative errors in air temperature measurements, *Water Resour. Res.*, *45*, W08431, doi:10.1029/2008WR007600.
Jelinek, C. (1869), *Anleitung zur Anstellung meteorologischer Beobachtungen und Sammlung von Hilfstafeln*, K.k. Hof- und Staatsdruckerei, 193 pp., Vienna.
Klein Tank, A., et al. (2002), Daily dataset of 20th-century surface air temperature and precipitation series for the European climate assessment, *Int. J. Climatol.*, *22*(12), 1441–1453, doi:10.1002/joc.773.
Kuglitsch, F., R. Auchmann, R. Bleisch, S. Brönnimann, O. Martius, and M. Stewart (2012), Break detection of annual Swiss temperature series, *J. Geophys. Res.*, *117*, D13105, doi:10.1029/2012JD017729.
Maugeri, M., L. Buffoni, B. Delmonte, and A. Fassina (2002), Daily Milan temperature and pressure series (1763–1998): Completing and homogenising the data, *Clim. Change*, *53*(1–3), 119–149, doi:10.1023/A:1014923027396.
Menne, M. J., I. Durre, R. S. Vose, B. E. Gleason, and T. G. Houston (2012), An overview of the global historical climatology network-daily database, *J. Atmos. Oceanic Technol.*, *29*(7), 897–910, doi:10.1175/JTECH-D-11-00103.1.
Mestre, O., C. Gruber, C. Prieur, H. Caussinus, and S. Jourdain (2011), SPLIDHOM: A method for homogenization of daily temperature observations, *J. Appl. Meteorol. Climatol.*, *50*(11), 2343–2358, doi:10.1175/2011JAMC2641.1.
Moberg, A., et al. (2006), Indices for daily temperature and precipitation extremes in Europe analyzed for the period 1901–2000, *J. Geophys. Res.*, *111*, D22106, doi:10.1029/2006JD007103.
Mueller, B., and S. I. Seneviratne (2012), Hot days induced by precipitation deficits at the global scale, *Proc. Natl. Acad. Sci. U.S.A.*, *109*(31), 12,398–12,403, doi:10.1073/pnas.1204330109.
Nemec, J., C. Gruber, B. Chimani, and I. Auer (2013), Trends in extreme temperature indices in Austria based on a new homogenised dataset, *Int. J. Climatol.*, *33*(6), 1538–1550, doi:10.1002/joc.3532.
Parker, D. (1994), Effects of changing exposure of thermometers at land stations, *Int. J. Climatol.*, *14*(1), 1–31, doi:10.1002/joc.3370140102.
Pattarone, G., and G. Alice (1925), *Osservatorio Geofisico Rosmini. Domodossola. Dati pluviometrici raccolti nel primo cinquantennio*, 6 pp., Tip. Porta, Domodossola, Italy.
Pinauda, F. (1914), *Nozioni di Meteorologia Ossolana, Ossia, Il clima dell'Ossola Superiore Desunto Dalle Osservazioni del Quarantennio 1872–1911*, 117 pp., Edizioni L. Alfieri, Domodossola, Italy.
Sen, P. K. (1968), Estimates of the regression coefficient based on Kendall's tau, *J. Am. Stat. Assoc.*, *63*(324), 1379–1389, doi:10.1080/01621459.1968.10480934.
Sneyers, R. (1990), On the statistical analysis of series of observations, *WMO Tech. Note 143*, Geneva, Switzerland.
Stocker, T., D. Qin, G.-K. Plattner, M. Tignor, S. K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley (Eds.) (2014), *Climate Change 2013: The Physical Science Basis*, Cambridge Univ. Press, Cambridge, U.K., and New York.
Sutton, R. T., and D. L. Hodson (2005), Atlantic ocean forcing of North American and European summer climate, *Science*, *309*(5731), 115–118, doi:10.1126/science.1109496.
Toreti, A., F. G. Kuglitsch, E. Xoplaki, J. Luterbacher, and H. Wanner (2010), A novel method for the homogenization of daily temperature series and its relevance for climate change analysis, *J. Clim.*, *23*(19), 5325–5331, doi:10.1175/2010JCLI3499.1.
Toreti, A., F. G. Kuglitsch, E. Xoplaki, and J. Luterbacher (2012), A novel approach for the detection of inhomogeneities affecting climate time series, *J. Appl. Meteorol. Climatol.*, *51*(2), 317–326, doi:10.1175/JAMC-D-10-05033.1.
Trewin, B. (2013), A daily homogenized temperature data set for Australia, *Int. J. Climatol.*, *33*(6), 1510–1529, doi:10.1002/joc.3530.
Van der Meulen, J., and T. Brandsma (2008), Thermometer screen intercomparison in De Bilt (The Netherlands), Part I: Understanding the weather-dependent temperature differences, *Int. J. Climatol.*, *28*(3), 371–388, doi:10.1002/joc.1531.
Venables, W., and B. Ripley (2002), *Modern Applied Statistics Using S*, Springer, New York.
Vose, R. S., D. R. Easterling, and B. Gleason (2005), Maximum and minimum temperature trends for the globe: An update through 2004, *Geophys. Res. Lett.*, *32*, L23822, doi:10.1029/2005GL024379.
Wacker, S., J. Gröbner, C. Zysset, L. Diener, P. Tzoumanikas, A. Kazantzidis, L. Vuilleumier, R. Stöckli, S. Nyeki, and N. Kämpfer (2015), Cloud observations in Switzerland using hemispherical sky cameras, *J. Geophys. Res. Atmos.*, *120*, 695–707, doi:10.1002/2014JD022643.
Wang, X. L., Q. H. Wen, and Y. Wu (2007), Penalized maximal t test for detecting undocumented mean change in climate data series, *J. Appl. Meteorol. Climatol.*, *46*(6), 916–931, doi:10.1175/JAM2504.1.
Weber, R. O., P. Talkner, I. Auer, R. Böhm, M. Gajić-Čapka, K. Zaninović, R. Brazdil, and P. Faško (1997), 20th-century changes of temperature in the mountain regions of central Europe, *Clim. Change*, *36*(3–4), 327–344, doi:10.1023/A:1005378702066.
Yan, Z., et al. (2002), Extreme temperature trends in Europe and China based on daily observations, *Clim. Change*, *53*, 355–392, doi:10.1007/978-94-010-0371-1_13.